A detailed physical map of the horse Y chromosome

Terje Raudsepp*†, Avni Santani*†, Barbara Wallner‡, Srinivas R. Kata§, Chengwei Ren¶, Hong-Bin Zhang¶, James E. Womack§, Loren C. Skow*, and Bhanu P. Chowdhary*‖**

Departments of *Veterinary Anatomy and Public Health, §Veterinary Pathobiology, ¶Soil and Crop Sciences and Crop Biotechnology Center, and ‖Animal Science, Texas A&M University, College Station, TX 77843; and ‡Institut für Tierzucht und Genetik, Veterinärmedizinische Universität Wien, A-1210 Vienna, Austria

Contributed by James E. Womack, April 28, 2004

We herein report a detailed physical map of the horse Y chromosome. The euchromatic region of the chromosome comprises ≈15 megabases (Mb) of the total 45- to 50-Mb size and lies in the distal one-third of the long arm, where the pseudoautosomal region (PAR) is located terminally. The rest of the chromosome is predominantly heterochromatic. Because of the unusual organization of the chromosome (common to all mammalian Y chromosomes), a number of approaches were used to crossvalidate the results. Analysis of the 5,000-rad horse × hamster radiation hybrid panel produced a map spanning 88 centirays with 8 genes and 15 sequence-tagged site (STS) markers. The map was verified by several fluorescence *in situ* hybridization approaches. Isolation of bacterial artificial chromosome (BAC) clones for the radiation hybrid-mapped markers, end sequencing of the BACs, STS development, and bidirectional chromosome walking yielded 109 markers (100 STS and 9 genes) contained in 73 BACs. STS content mapping grouped the BACs into seven physically ordered contigs (of which one is predominantly ampliconic) that were verified by metaphase-, interphase-, and fiber-fluorescence *in situ* hybridization and also BAC fingerprinting. The map spans almost the entire euchromatic region of the chromosome, of which 20–25% (≈4 Mb) is covered by isolated BACs. The map is presently the most informative among Y chromosome maps in domesticated species, third only to the human and mouse maps. The foundation laid through the map will be critical in obtaining complete sequence of the euchromatic region of the horse Y chromosome, with an aim to identify Y specific factors governing male infertility and phenotypic sex variation.

gene map

Significant progress has been made during recent years to analyze the equine genome. Presently, low- to medium-resolution synteny, genetic linkage, and cytogenetic maps are available for all horse autosomes and the X chromosome. More recently, analysis of the 5,000-rad horse × hamster radiation hybrid (RH) panel (1) provided the first-generation RH and comparative map for the horse and facilitated the construction of ≈1-Mb-resolution physically ordered maps for some of the horse chromosomes [e.g., ECA17 (2), ECAX (3); ECA22 (A. L. Gustafson, T.R., M. L. Wagner, J. R. Mickelson, L.C.S., and B.P.C., unpublished work]. Despite this progress, one chromosome that has remained devoid of mapped markers is the equine (*Equus caballus*) Y chromosome (ECAY). Very limited effort has been made until now to develop a gene map for this chromosome or to understand its structure and organization (4–6).

The mammalian Y chromosome has long been considered to have gradually degenerated during evolution and to have lost most of the functional genes, preserving only those involved in sex determination. Consequently, few efforts have been directed to study the structure/organization and gene content of the Y chromosome in the majority of the species that presently have medium- to high-resolution gene maps. The Y chromosome has been inappropriately ignored as a "barren wasteland" and has frequently been referred to as mainly containing "junk" (7). However, recent studies in humans clearly indicate the presence of at least 27 protein-coding genes/gene families, of which ≈10 are specifically expressed in testes (8), suggesting that these sequences may function in sperm development and fertility. Additionally, several studies implicate the Y chromosome in male germ cell development and maintenance (9, 10), certain types of cancers (11), and other important biological functions (10, 12). Several of these traits are also of considerable significance in livestock, pet, and companion animals, for which the understanding of the role of Y chromosome genes is especially lacking.

Among mammals, detailed information on the organization and gene content of the Y chromosome is presently available in humans (8, 13). Overall, the chromosome is largely heterochromatic with extensive tandem repeats. The euchromatic region contains fewer protein-coding genes than the rest of the genome. The situation is further complicated by the presence of large palindromes and dispersed multicopy gene families (7, 8, 13, 14). On the whole, this bizarre organization is the main reason why the generation of a gene map for the human Y chromosome was extremely tedious and time consuming compared to autosomes and the X chromosome (8, 13). The impediment is further reflected in the fact that the draft and complete sequence of the human genome did not contain comprehensive information on the Y chromosome. Very recently, however, using a range of approaches, the male specific region of the human Y was completely sequenced (8). Therefore, it comes as no surprise that mapping information is very limited for the Y chromosomes in other mammalian species, e.g., ape (15), mouse (16, 17), pig (18), cat (19), cattle (20), and dog (21).

Stallion fertility is of prime importance to the equine industry, where stud fees can be enormous. Presently, very little is known about the molecular mechanisms regulating stallion fertility. Even less is known about the genes governing them. To complicate matters further, virtually nothing is known about the genetic component of fertility that is male specific and is located on the Y chromosome. The latter is evident from the current sparse map of the horse Y chromosome (ECAY). Thus far, only three genes (*SRY*, *ZFY*, and *STS*) are physically assigned to ECAY by fluorescence *in situ* hybridization (FISH) (6) and analysis of somatic hybrid cell panel (5). Except for identification of *SRY* as the sex determination factor (4, 22), practically no information is available about the association of Y chromosome sequences with stallion infertility. This study reports the generation of a detailed physical map of the ECAY obtained by using a range of mapping approaches, including RH analysis; metaphase-, interphase-, and DNA fiber-FISH; chromosome walking; bacterial artificial chromosome (BAC) contig generation; and

GENETICS

Abbreviations: RH, radiation hybrid; ECAY, *Equus caballus* Y chromosome; FISH, fluorescence *in situ* hybridization; BAC, bacterial artificial chromosome; Mb, megabases; PAR, pseudoautosomal region; STS, sequence-tagged site.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. can be found in Table 1, which is published as supporting information on the PNAS web site).

†T.R. and A.S. contributed equally to this work.

**To whom correspondence should be addressed. E-mail: bchowdhary@cvm.tamu.edu.

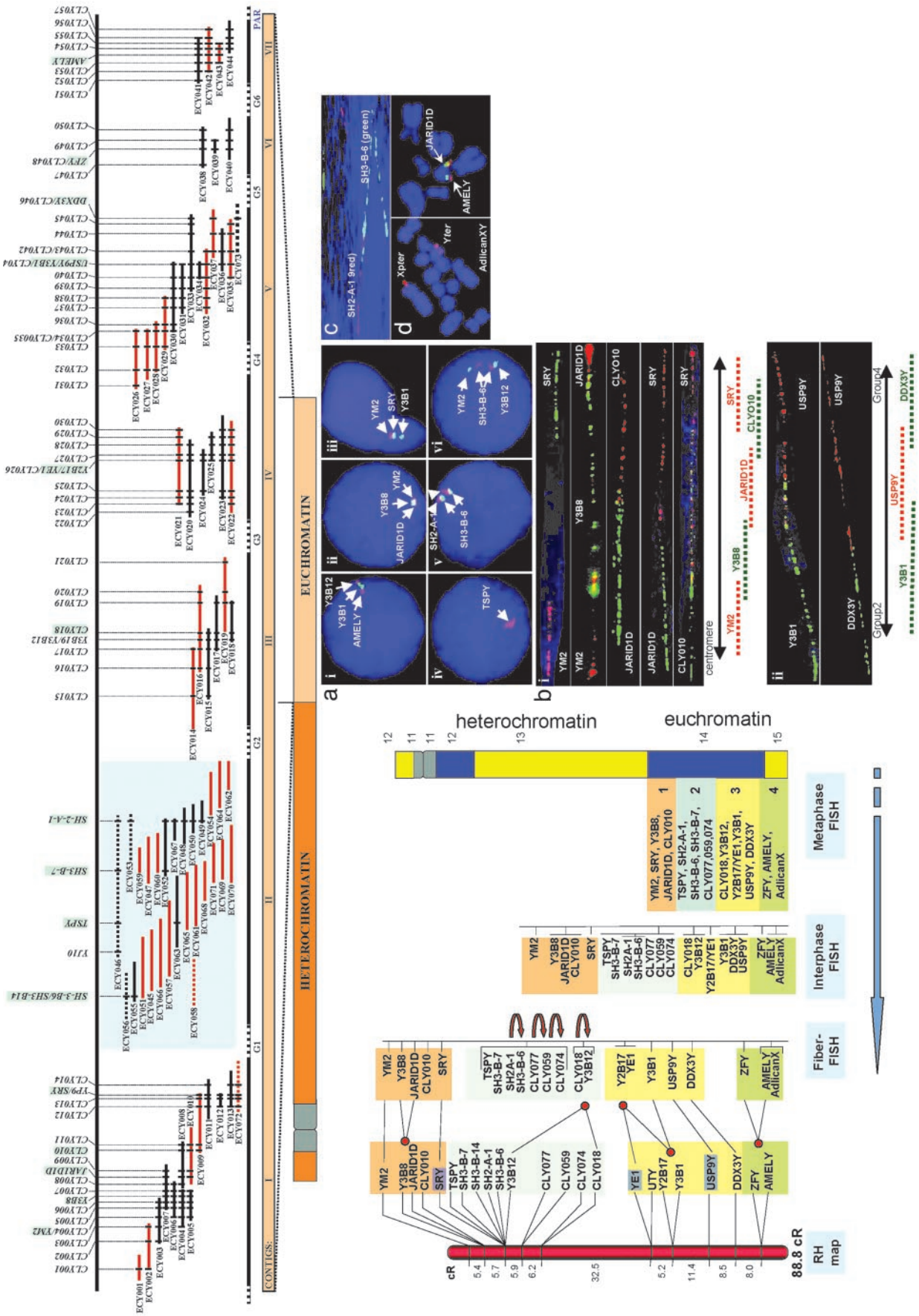

Fig. 1. (*Legend appears at the bottom of the opposite page.*)

BAC fingerprinting. An array of BAC contigs thus developed over the euchromatic region of the ECAY marks the first step toward finding Y specific genes/sequences involved in male fertility in horses.

Materials and Methods

Primer Design, PCR Optimization, and Sequence Verification. Primer pairs were developed from 15 Y specific horse sequence-tagged site (STS) and five partial cDNA sequences representing functional genes [www.ncbi.nlm.nih.gov (23, 24)]. Additionally, three pairs of heterologous primers were designed by using published Y specific sequences for human *USP9Y* and *DDX3Y* and porcine *UTY*. Last, 57 end sequences from equine BAC clones isolated in this study for bidirectional chromosome walking and BAC contig construction (see below) were also used to generate primer sets, following REPEATMASKER analysis (http://repeatmasker.genome.washington.edu) and BLAST comparisons (www.ncbi.nlm.nih.gov/BLAST). All primers were designed by using PRIMER3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) and optimized on male horse, female horse, and hamster DNA (for RH mapping, see below) so that only male horse-specific PCR product was obtained. PCR products amplified with the three heterologous primers were sequenced as described elsewhere (2) to confirm their identity. Detailed information on markers, primers, PCR conditions, etc., is summarized in Table 1.

BAC Library Screening and End Sequencing. Optimized male-specific primers were used primarily to screen by PCR Children's Hospital Oakland Research Institute 241 (http://bacpac.chori.org/equine241.htm) and Texas A&M University equine BAC libraries (http://hbz7.tamu.edu/homelinks/bac_est/bac.htm), and to a lesser extent the Institut National de la Recherche Agronomique-complemented BAC library (25). Initial screening was done by using primers derived from 15 sequence-tagged sites and eight genes (see above). Subsequent screening was done by using primer pairs obtained from end sequences of BACs isolated in this study (see Table 1 for details). The screening proceeded from superpools to plate pools and then from individual plates to specific clones. DNA extraction was carried out by using a Qiagen midi-prep kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. BAC end sequencing was carried out by using the standard T7, SP6, and M13 reverse primers. Dye terminator sequencing reactions (total volume 10 μl) were set up and performed in a Gene Amp (Applied Biosystems) PCR system 9700 by using previously described protocols (26). Reaction products were purified by using spin columns (Spin-50, BioMax, Odenton, MD) and loaded on ABI 3100 automated capillary sequencers (PE Applied Biosystems) for analysis.

BAC Fingerprinting and Contig Assembly. BAC DNA was isolated and fingerprinted according to ref. 27. Briefly, BAC DNA was isolated, purified, and double-digested with *Hin*dIII and *Hae*III. The *Hin*dIII fragments were end-labeled with [^{32}P]dATP by using reverse transcriptase at 37°C for 2 h and then subjected to electrophoresis on 3.5% (wt/vol) polyacrylamide DNA sequencing gels at 90 W for 100 min. The gels were dried and autoradiographed. Fingerprint editing and contig assembly were conducted by scanning autoradiographs into digital image files by using a UMAX (Dallas) Mirage D-16L scanner and were edited by using IMAGE 3.10b software (28). Because of the lower resolution of the higher-molecular-weight bands at the top of the gels, only the bands ranging from 58 to 773 bp were used for contig assembly. Vector bands were removed manually from the data files. All digitized band data were standardized against the λ-DNA/*Sau*3AI marker and converted from base pairs into migration rates. Contigs were automatically assembled by FPC PROGRAM Version 6 (29). The tolerance was fixed at 2, and the cutoff value was set to 1e-06 by a series of tests.

RH Mapping. Primers for eight genes and 15 STS markers were typed in duplicate on the 5,000-rad horse $\times$ hamster RH panel (1). PCR products were visualized on 2% agarose gels. Scoring of positive clones and RH data analysis were performed as described in detail earlier (1, 30).

FISH. Peripheral blood from a male horse (Reg #1219, Texas A&M University) with known fertility was used for (*i*) standard short-term pokeweed (Sigma) stimulated lymphocyte cultures to obtain interphase and metaphase chromosome preparations and (*ii*) agarose-embedded DNA plugs to obtain DNA fibers (31, 32). DNA from individual BAC clones was labeled with biotin-16-dUTP and/or digoxigenin-11-dUTP by nick translation using Biotin- or DIG-Nick Translation Mix (Roche Molecular Biochemicals). Hybridization to metaphase chromosomes or mechanically stretched DNA fibers was performed as described elsewhere in detail by us (30, 32–34). A minimum of 30 metaphase spreads and $\approx$50 interphase cells or DNA fiber hybridizations were captured and analyzed for each experiment by using a Zeiss Axioplan2 fluorescent microscope equipped with CYTOVISION/GENUS application software, Version 2.7 (Applied Imaging, Santa Clara, CA).

Results

BAC Library Screening and End Sequencing. Screening the Texas A&M University, CHORI-241, and Institut National de la Recherche Agronomique horse genomic BAC libraries with the initial set of 23 markers (8 genes and 15 STS) provided a total of 37 BAC clones representing 20 of the markers. No positive clones were found for *UTY, YA16*, and *YH12*. Direct end sequencing of each of the clones provided nucleotide sequence information averaging 928 bp (range, 300–1,400 bp), which was used to develop 60 previously uncharacterized unique STSs. The remaining sequences were predominantly long interspersed repeats and therefore could not be used for further analysis. On the basis of the marker content, the BACs were manually grouped into nine contigs. This foundation assembly was subsequently used for bidirectional chromosome walking to obtain adjacent overlapping clones.

Bidirectional Chromosome Walking. Screening the BAC libraries with recently developed STS markers provided additional depth to the foundation contigs and also added few BACs to some of the termini. Subsequently, outbound end sequences of the "terminal" BACs (BACs lying at the ends of a contig) were used

Fig. 1. (*On the opposite page.*) A map of the euchromatic region of ECAY. (*Top*) A contig map of 73 BACs over the euchromatic region of ECAY. Top row represents STS loci and genes used to generate the seven contigs (I–VII). Loci shaded blue are also present in the RH map. Contig II (in blue rectangle) shows BACs ordered only by fingerprinting. G1–G6 represent gaps in the contigs. (*Bottom Left*) RH and cytogenetic FISH map of the ECAY. The RH map is correlated with maps obtained by various FISH approaches through the four color codes for groups of loci (1–4 beside Y chromosome ideogram). Fiber and interphase FISH resolved order for several markers. Semicircular arrows in the fiber-FISH map indicate partial overlaps between copies of some of the markers. (*Bottom Right*) (*a i–vi*) interphase FISH to resolve order of markers. (*b*) Two examples whereby fiber FISH resolved order of loci, as depicted in schematic drawings below each set of experiments. (*c*) Hybridization with two BACs containing multicopy sequences (*d*) (*Left*) Hybridization of *Adlican* on both the X and Y chromosomes; (*Right*) cohybridization showing the order of *JARID1D* (green) and *AMELY* (red).

for targeted expansion. Several rounds of walking/screening attempts provided 36 previously uncharacterized BACs at or close to the terminal ends. Few of these BACs facilitated closure of two of the gaps, i.e., between *SRY-JARID1D* and *YM2-JARID1D* (see Fig. 1). End sequencing of these BACs yielded 43 additional STS. Thus, the entire experiment of end sequencing and chromosome walking provided a total of 103 STS. However, 32 of the STS amplified a similar-size DNA band in the control female in addition to the male, whereas the remaining 71 were male specific. The average size of the STS was 186 bp (range 100–495 bp).

STS Content Mapping and Contig Assembly. Analysis by PCR of the 73 BACs with 23 initial markers and 103 newly developed STS provided 7 groups of overlapping BAC contigs. Six of the contigs (I and III–VII; Fig. 1) were comprised of 3–13 BACs. Contig II, referred to as the putative "multicopy" region (discussed later), had 27 BACs. Within each of the former six contigs, the likely order, orientation, and overlap of BACs were determined on the basis of the presence or absence of specific markers. A summary of results showing individual contigs (I and III—VII, oriented centromere to telomere) and all markers tested for content mapping of BACs are presented in Fig. 1. BACs in the "multicopy group" (contig II; light-blue shaded region, Fig. 1) were ordered by using a fingerprinting approach. The physical order of the contigs was derived by using a combination of RH and interphase-/fiber-FISH approaches and is described in pertinent sections.

Restriction digestion and fingerprint analysis of 67 BACs (6 BACs did not grow) gave a total of 2,466 unique bands at a tolerance of 2 and a cutoff at 1e-06 and allowed the BACs to be assembled into seven contigs. One clone remained as a singleton. Overall results show that the fingerprint map (not presented) is in close agreement with the contigs obtained by STS content mapping.

RH Mapping. Genotyping of the initial 23 markers (8 genes and 15 STSs) on the 5,000-rad horse × hamster RH panel provided a map that spanned 88 centirays and covered almost the entire euchromatic region of the ECAY (Fig. 1). The average retention frequency of the markers was 11%. Two-point linkage analysis of typing data resulted in a single linkage group at a logarithm of odds score of 7 (2PT-RHMAP; ref. 35). Anchoring the RH map to chromosome by FISH enabled us to orient the map and indicated that *YM2*, *JARID1D*, *SRY*, and *TSPY* are proximally located within the euchromatic region, whereas USP9Y, *DDX3Y*, *ZFY*, and *AMELY* are distal. At seven of the map locations, two or more markers clustered at the same position (Fig. 1), but the relative order of these loci could not be determined. Other approaches described below were used to resolve physical order for these loci.

FISH Mapping. ***Metaphase FISH.*** Of the 114 BAC clones isolated in this study, 67 hybridized specifically to the Y chromosome, whereas four hybridized to both X and Y. Six of the clones hybridized both to Y and an autosome. Of the remaining clones, 22 showed signals on only autosomes, two only on the X chromosome, and 13 did not give a specific signal. These hybridizations helped to verify the origin of individual clones. A total of 22 clones were selected and used for cytogenetic localization of the RH mapped markers (see Table 1). Markers *Y2B17/YE1* and *SH3-B-14/SH3-B*$_6$ were present in the same BACs. *AdlicanXY* containing BAC was taken from the *Equus caballus* X chromosome (3). Single and double-color metaphase FISH with various combinations of markers helped to divide the 22 loci into four major physically ordered groups in the euchromatic region of ECAY (groups 1–4; see Fig. 1). For example, cohybridization of *JARID1D* (group 1) and *AMELY* (group 4) clearly showed a noteworthy gap between them and also elaborated their proximal and distal orientation on the chromosome (Fig. 1*d Right*). Due to the diminutive size of ECAY euchromatin and the resolution limit of ≈2–3 Mb for metaphase FISH (36), it was not possible to resolve the order of markers within the groups.

Interphase FISH. Double-color hybridizations on interphase nuclei with labeled probes used in different combinations confirmed the deduced physical organization of the four groups and allowed us to resolve the physical order of the majority of the loci within each group, in particular the distal groups 3 and 4 (Fig. 1*a*). For example, *Y3B12* was found to be the most proximal marker in group 3. In group 4, *AMELY* was the most terminal marker (Fig. 1*ai*), and *ZFY* was distinctly proximal to *AMELY*. However, BACs for *Y3B1*, *USP9Y*, and *DDX3Y* in group 3 and for *AMELY* and *AdlicanXY* in group 4 gave strongly overlapping signals on interphase chromosomes, hence their order remained ambiguous. Similarly, signals for markers in the most proximal group (*YM2*, *Y3B8*, *JARID1D*, and *SRY*) mostly overlapped (Fig. 1*aii*). We could, nevertheless, show that the end markers of this group were *YM2* and *SRY* by cohybridization of BACs containing these two loci with markers from other groups. The results revealed that *YM2* is the most proximal and *SRY* is the most distal within this group (Fig. 1*aiii*).

It is remarkable that all markers in group 2 (Fig. 1) showed two or more hybridization signals in interphase nuclei, indicating the presence of more than one copy of these sequences on ECAY. The BAC containing *TSPY* showed the strongest hybridization signal (Fig. 1*aiv*). Other markers gave ≈2–10 signals in interphase cells. Therefore, cohybridization of markers from this region could not resolve their relative order (Fig. 1*av*). Nonetheless, it appears that signals from these BACs are regionally clustered and the majority of the multiple copies are located between group 1 (e.g., *SRY*) and group 3 (e.g., *Y3B12*; Fig. 1*avi*). However, the presence of these copies in other regions of the ECAY euchromatin cannot be excluded.

Fiber FISH. Hybridizations with combinations of two or three probes on mechanically stretched DNA fibers from the most proximal cluster (group 1) enabled precise determination of the order of five markers as: prox-*YM2-Y3B8-JARID1D-CLY010-SRY*-dist (Fig. 1*bi*). Similarly, the relative order of markers in group 3 was resolved as: prox-*Y3B1-USP9Y-DDX3Y*-dist (Fig. 1*bii*). However, BAC clones for markers in the terminal group 4 (*AdlicanX* and *AMELY*) showed neither proximity nor overlaps on DNA fibers. Hence their relative order could not be resolved (Fig. 1).

As expected, ordering markers in the multicopy cluster was difficult. Nonetheless, we could confirm and refine interphase results concerning close proximity of *SH2-A-1* and *SH3-B*$_6$, showing that at least one copy of *SH2-A-1* partially overlapped with at least one copy of *SH3-B*$_6$ on fibers (Fig. 1*c*). Fiber FISH with *TSPY* confirmed our interphase observations that the gene has several copies on ECAY. Further, cohybridizations indicated that some copies of marker pairs *CLY077-CLY059* and *CLY059-CLY07*, respectively, either overlap or are in close proximity to each other. However, due to the presence of numerous nonoverlapping copies, the physical order of these STSs could not be resolved (Fig. 1).

Discussion

Overview of Map and Markers. The study provides a detailed physical map of the ECAY. RH analysis and a range of FISH mapping approaches provided 23 markers on a physically ordered foundation map for the Y chromosome. The markers, together with the additional set developed in this study by us (total 113; 104 STS and 9 genes) allowed us to generate an elaborate contig map containing seven groups of 73 overlapping BAC contigs spread over the euchromatic region of the ECAY. Overall, the study represents a major development, because gene maps are presently available for only equine autosomes and the X chromosome (30, 37).

Information on the organization of the Y chromosome in mammals other than human and mouse is sparse. At present, a

total of 62 loci (primarily microsatellites) are known to have been mapped in cattle (20, 38); 8 genes in the cat (19); and 10 loci in the dog, of which *SRY* is the only gene (21); and 10 genes in the pig (18). Even in rats, only nine genes have been provisionally assigned to the Y chromosome (http://ratmap.gen.gu.se/SearchList.html). Considering this, the ECAY map presented in this study signifies a substantial advancement over corresponding maps in other species, making it third in mapped loci after human and mouse Y maps.

Euchromatic Region and PAR on ECAY. ECAY most likely spans ≈45–50 Mb. This is estimated from the relative length of 1.62% (39) against a 3,000-Mb size of the equine genome. Like other mammalian species, a considerable part of ECAY is heterochromatic. The euchromatic part lies toward the distal one-third of the chromosome and is expected to be ≈15 Mb. Developing 7 contigs with 73 Y-specific BACs provides a coverage of the ECAY euchromatic region beyond that reported in other domesticated species. The BACs spanning the minimum tiling path for each of the contigs indicate a cumulative coverage of ≈4 Mb that corresponds to almost 20–25% of the euchromatic region.

Synaptonemal complex analysis suggests the equine PAR on Y to be located on the terminal part of the long arm of Y chromosome (40). In several primates, cattle, pig, and horse, the *AMEL* locus spans the PAR boundary (41). The location of *AMEL* in the telomeric part of our map indicates the likely location of ECAY-PAR toward the distal end of the euchromatic region. Mapping of more genes will help to elucidate this region.

Map Alignment and Contig Development. Markers common to the contigs and the RH/FISH map (highlighted in Fig. 1) were critical in aligning the two maps and in deducing the physical order of the seven contigs within the euchromatic region. In combination with STS content mapping, the markers, in effect, provided an ordered map of 77 markers spanning from near the proximal to near the terminal end of the euchromatic region. The order of the markers, however, is preliminary, because other than in contigs I and V, the proposed centromere⇒telomere orientation of the loci could be reversed. However, within individual contigs, the order was cross-verified by STS content mapping and/or fiber FISH. It is noteworthy that the contigs are fairly uniformly distributed across the euchromatic region. Hence, despite the six gaps between the contigs, the map provides a valuable platform for further chromosome walking and contig expansion.

Multiple Approaches to Develop a Y Map. A combination of mapping approaches was simultaneously used to develop the ECAY map. The unusual organization of this chromosome necessitated cross-verification and confirmation of results, because observations from a single approach were inconclusive. For example, the RH map required verification and refinement with a range of FISH mapping techniques, because at seven of the locations, marker order was unclear (see RH map, Fig. 1). This is not unexpected, considering that within a small genomic region (15–20 Mb for euchromatic region of Y) the resolution limit of mapping in a 5,000-rad RH panel might easily be saturated with ≈20 markers due to insufficient number of breaks in the haploid Y component. In such instances, interphase and fiber FISH proved instrumental in resolving the order of 17 of the 23 mapped loci, thus refining and validating RH and metaphase-FISH data. The approaches were also critical in confirming the *cen⇒tel* orientation of the seven BAC contigs. Further, STS content mapping permitted verification of the results. For example, the order of loci *YM2-Y3B8-JARID1D-CLY10-SRY* deduced from the FISH map (Fig. 1) was supported independently by STS content mapping (contig I, Fig. 1). Similarly, the STS content mapping order of *Y3B1-USP9Y-DDX3Y* (contig V, Fig. 1) corresponded to that obtained by fiber FISH (Fig. 1*b*). Last, fingerprint analysis of the BACs provided vital supplementary confirmation regarding the grouping and ordering of BACs/markers.

Putative Ampliconic Region(s) on ECAY. It was noteworthy that aligning and ordering of BACs in contig II (shaded block in Fig. 1) by STS content mapping, interphase FISH, or fiber FISH were virtually impossible. (Fig. 1). Eight of the STS from this region showed PCR amplification on BAC templates from contigs I, V, and VI. Further, 11 markers from contigs I, IV, and V showed PCR amplification on BACs from contig II. Thus contig II bore all of the hallmarks of a "multicopy or ampliconic region," with contigs I and V as prospective sites for additional ampliconic sequences. Therefore, restriction fingerprinting was the only way to deduce a putative order of clones in this region. Despite this, the physical placement of the contig in relation to other contigs is definitive based on (*i*) the location of five of the six markers from this contig in the RH map and (*ii*) interphase and fiber-FISH orientation of BACs from this region, in relation to BACs from contigs I and III.

Multicopy or *ampliconic* regions have been typically found in human (8), primate (14), mouse (42), and cattle (20). In humans, the ampliconic region cumulatively spans ≈10 Mb and comprises 24 gene families and 9 single-copy units (8). At present, it is difficult to ascertain the number, distribution, and size of ampliconic region(s) in the horse. Nonetheless, the contig II (Fig. 1) assembled by fingerprinting indicates the region to be ≈1.2 Mb. Expansion of this contig and identification of other ampliconic regions will help to provide an accurate cumulative estimate.

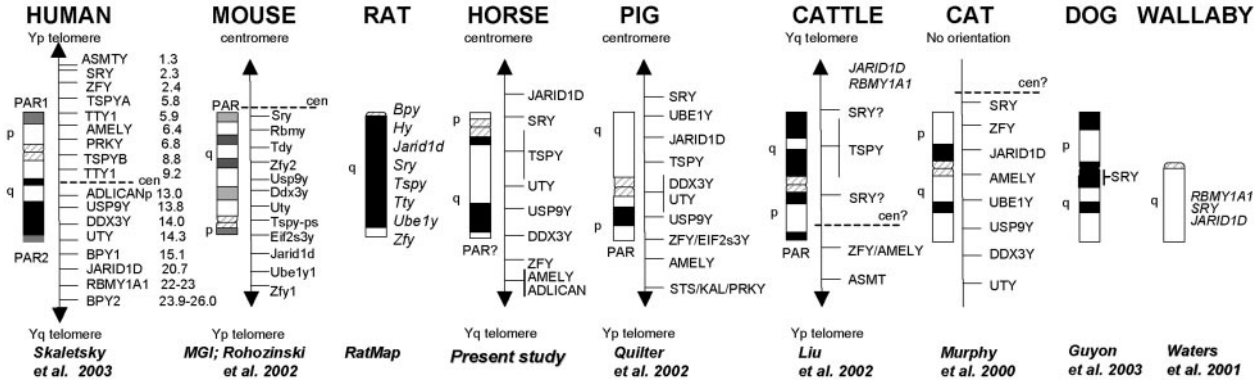

Fig. 2. An overview of the comparative status of Y chromosome maps in several mammalian species. Where available, a physical order of loci is provided. In humans, only those loci are shown that are present on the Y chromosome map of other species. *UBE1*, *Eif2s3Y*, *HY*, and *TDY* are not found in humans. Loci shown in italics (rat, cattle, and wallaby) are not ordered.

Comparative Map. Earlier crossspecies comparisons of Y chromosome by Zoo FISH suggested lack of homology across evolutionarily distantly related species (e.g., humans vs. mouse, cattle pig, dog, or horse; ref. 43) but the presence of homology across closely related species, e.g., within primates (44), bovids (unpublished results), and equids (45). Morphologically, the Y chromosome is almost meta-/submetacentric in human, cattle, cat, dog, and pig. However, it is almost acrocentric in mouse, rat, and horse. Not more than 10 Y specific genes have been mapped until now in nonhuman/mouse species. A comparative overview of the physical order of these genes in different species (Fig. 2) indicates that ECAY most closely resembles the porcine counterpart. In both species, *ZFY* and *AMELY* are telomeric/distal (like in cattle), whereas *SRY* and *JARID1D* are centromeric/proximal (as observed in cat and mouse but not in cattle and humans). Additionally, *TSPY*, *UTY*, *DDX3Y*, and *USP9Y* are clustered together between the two terminal groups in both species, and their order is also essentially conserved. Incidentally, these four loci are also clustered together in human and mouse. Other than this, no clear conservation in gene order was observed for genes mapped in different species. It is noteworthy that *DAZ*, which maps to the Y chromosome in humans, is autosomal in the horse and is located on ECA16q22.3 (unpublished results).

Conclusion

As in most mammalian species, ECAY is among the smallest chromosomes in the karyotype and is atypical and unusual in organization and gene content compared to other chromosomes. Usual mapping strategies used to rapidly develop gene maps and BAC contigs on autosomes and the X chromosome are of little consequence for the Y. Hence developing a map of this chromosome, obtaining genomic clones representative of the entire chromosome, developing contigs, and acquiring sequence information pose special challenges that are specific only to this chromosome. A combination of mapping approaches required to generate and closely verify the first-generation physical map of ECAY is a testament to this inherent problem. This map will serve as an important foundation for expanded studies aimed at developing a minimum tiling path of BACs over the euchromatic region that could finally be used for obtaining a complete sequence of the region. The findings may be of specific significance in initiating studies aimed at identifying Y chromosome-related factors associated with spermatogenesis failures, stallion infertility, and regulation of the phenotypic sex. These investigations may also add to current research on similar lines in humans.

This project was funded by grants from the American Quarter Horse Association, the Texas Higher Education Board [ARP 010366-0162-2001 (to B.P.C.) and ATP 000517-0306-2003 (to B.P.C. and J.E.W.)], National Research Initiative Competitive Grants Program/U.S. Department of Agriculture Grant 2003-03687 (to B.P.C.), the Texas Equine Research Foundation (to B.P.C. and L.C.S.), the Link Endowment (to B.P.C. and L.C.S.), and the Dorothy Russell Havemeyer Foundation. Additional support was available from the U.S. Department of Agriculture–National Research Support Projects-8 Coordinators Fund.

1. Chowdhary, B. P., Raudsepp, T., Honeycutt, D., Owens, E. K., Piumi, F., Guérin, G., Matise, T. C., Kata, S. R., Womack, J. E. & Skow, L. C. (2002) *Mamm. Genome* **13,** 89–94.
2. Lee, E.-J., Raudsepp, T., Kata, S. R., Adelson, D., Womack, J. E., Skow, L. C. & Chowdhary, B. P. (2004) *Genomics* **83,** 203–215.
3. Raudsepp, T., Lee, E. J., Kata, S. R., Brinkmeyer, C., Mickelson, J. R., Skow, L. C., Womack, J. E. & Chowdhary, B. P. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 2386–2391.
4. Hasegawa, T., Sato, F., Ishida, N., Fukushima, Y. & Mukoyama, H. (2000) *J. Vet. Med. Sci.* **62,** 1109–1110.
5. Shiue, Y.-L., Millon, L. V., Skow, L. C., Honeycutt, D., Murray, J. D. & Bowling, A. T. (2000) *Chromosome Res.* **8,** 45–55.
6. Hirota, K., Piumi, F., Sato, F., Ishida, N., Guerin, G., Miura, N. & Hasegawa, T. (2001) *Anim. Genet.* **32,** 326–327.
7. Lahn, B. T. & Page, D. C. (1997) *Science* **278,** 675–680.
8. Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., Repping, S., Pyntikova, T., Ali, J., Bieri, T., *et al.* (2003) *Nature* **423,** 825–837.
9. Sun, C., Skaletsky, H., Birren, B., Devon, K., Tang, Z., Silber, S., Oates, R. & Page, D. C. (1999) *Nat. Genet.* **23,** 429–432.
10. Quintana-Murci, L. & Fellous, M. (2001) *J. Biomed. Biotechnol.* **1,** 18–24.
11. Lau, Y. F. & Zhang, J. (2000) *Mol. Carcinog.* **27,** 308–321.
12. Ali, S. & Hasnain, S. E. (2003) *Gene* **321,** 25–37.
13. Tilford, C. A., Kuroda-Kawaguchi, T., Skaletsky, H., Rozen, S., Brown, L. G., Rosenberg, M., McPherson, J. D., Wylie, K., Sekhon, M., Kucaba, T. A., *et al.* (2001) *Nature* **409,** 943–945.
14. Charlesworth, B. (2003) *Genome Biol.* **4,** 226.
15. Rozen, S., Skaletsky, H., Marszalek, J. D., Minx, P. J., Cordum, H. S., Waterston, R. H., Wilson, R. K. & Page, D. C. (2003) *Nature* **423,** 873–876.
16. Mazeyrat, S., Saut, N., Sargent, C. A., Grimmond, S., Longepied, G., Ehrmann, I. E., Ellis, P. S., Greenfield, A., Affara, N. A. & Mitchell, M. J. (1998) *Hum. Mol. Genet.* **7,** 1713–1724.
17. Affara, N. A. & Mitchell, M. J. (2000) *J. Endocrinol. Invest.* **23,** 630–645.
18. Quilter, C. R., Blott, S. C., Mileham, A. J., Affara, N. A., Sargent, C. A. & Griffin, D. K. (2002) *Mamm. Genome* **13,** 588–594.
19. Murphy, W. J., Sun, S., Chen, Z. Q., Pecon-Slattery, J. & O'Brien, S. J. (1999) *Genome Res.* **9,** 1223–1230.
20. Liu, W. S., Mariani, P., Beattie, C. W., Alexander, L. J. & Ponce De Leon, F. A. (2002) *Mamm. Genome* **13,** 320–326.
21. Guyon, R., Lorentzen, T. D., Hitte, C., Kim, L., Cadieu, E., Parker, H. G., Quignon, P., Lowe, J. K., Renier, C., Gelfenbeyn, B., *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100,** 5296–5301.
22. Hasegawa, T., Ishida, M., Harigaya, T., Sato, F., Ishida, N. & Mukoyama, H. (1999) *J. Vet. Med. Sci.* **61,** 97–100.
23. Wallner, B., Brem, G., Muller, M. & Achmann, R. (2003) *Anim. Genet.* **34,** 453–456.
24. Wallner, B., Piumi, F., Brem, G., Muller, M. & Achmann, R. (2004) *J. Hered.* **95,** 158–164.
25. Milenkovic, D., Oustry-Vaiman, A., Lear, T. L., Billault, A., Mariat, D., Piumi, F., Schibler, L., Cribiu, E. & Guerin, G. (2002) *Mamm. Genome* **13,** 524–534.
26. Gustafson, A. L., Tallmadge, R. L., Ramlachan, N., Miller, D., Bird, H., Antczak, D. F., Raudsepp, T., Chowdhary, B. P. & Skow, L. C. (2003) *Cytogenet. Genome Res.* **102,** 189–195.
27. Ren, C., Lee, M. K., Yan, B., Ding, K., Cox, B., Romanov, M. N., Price, J. A., Dodgson, J. B. & Zhang, H. B. (2003) *Genome Res.* **13,** 2754–2758.
28. Sulston, J., Mallett, F., Durbin, R. & Horsnell, T. (1989) *Comput. Appl. Biosci.* **5,** 101–106.
29. Soderlund, C., Humphray, S., Dunham, A. & French, L. (2000) *Genome Res.* **10,** 1772–1787.
30. Chowdhary, B. P., Raudsepp, T., Kata, S. R., Goh, G., Millon, L. V., Allan, V., Piumi, F., Guérin, G., Swinburne, J., Binns, M., *et al.* (2003) *Genome Res.* **13,** 742–751.
31. Heiskanen, M., Karhu, R., Hellsten, E., Peltonen, L., Kallioniemi, O. P. & Palotie, A. (1994) *BioTechniques* **17,** 928–9, 932–933.
32. Sjoberg, A., Peelman, L. J. & Chowdhary, B. P. (1997) *Chromosome Res.* **5,** 247–253.
33. Raudsepp, T., Kijas, J., Godard, S., Guérin, G., Andersson, L. & Chowdhary, B. P. (1999) *Mamm. Genome* **10,** 277–282.
34. Liu, W. S., Soldatov, N. M., Gustavsson, I. & Chowdhary, B. P. (1998) *Hereditas* **129,** 169–175.
35. Boehnke, M. (1992) *Ann. Med.* **24,** 383–386.
36. Trask, B. J., Allen, S., Massa, H., Fertitta, A., Sachs, R., van den Engh, G. & Wu, M. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58,** 767–775.
37. Chowdhary, B. P. & Bailey, E. (2003) *Cytogenet. Genome Res.* **102,** 184–188.
38. Liu, W. S., Beattie, C. W. & Ponce de Leon, F. A. (2003) *Cytogenet. Genome Res.* **102,** 53–58.
39. Hansen, K. M. (1984) in *Proceedings of the 6th European Colloquium on Cytogenetics of Domestic Animals* (Zurich), pp. 165–171.
40. Power, M. M., Gustavsson, I., Switonski, M. & Ploen, L. (1992) *Cytogenet. Cell Genet.* **61,** 202–207.
41. Iwase, M., Satta, Y., Hirai, Y., Hirai, H., Imai, H. & Takahata, N. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 5258–5263.
42. Touré, A., Grigoriev, V., Mahadevaiah, S. K., Rattigan, A., Ojarikre, O. A. & Burgoyne, P. S. (2004) *Genomics* **83,** 140–147.
43. Chowdhary, B. P., Raudsepp, T., Fronicke, L. & Scherthan, H. (1998) *Genome Res.* **8,** 577–589.
44. Koehler, U., Bigoni, F., Wienberg, J. & Stanyon, R. (1995) *Genomics* **30,** 287–292.
45. Raudsepp, T. & Chowdhary, B. P. (1999) *Chromosome Res.* **7,** 103–114.